# PARTIALLY OCCLUDED OBJECT DETECTION IN VIDEO SEQUENCES


Submitted by

**Supriya Agrawal**

SAP ID: 70821016001


*for the award of the Degree of*

**DOCTOR OF PHILOSOPHY**

**(Engineering)**

Under the guidance of

**Dr. Prachi Natu**

**MUKESH PATEL SCHOOL OF TECHNOLOGY MANAGEMENT & ENGINEERING**

**SVKM's NARSEE MONJEE INSTITUTE OF MANAGEMENT STUDIES**

(Declared as Deemed-to-be University Under Section 3 of the UGC Act, 1956)

V. L. Mehta Road, Vile Parle (West)

MUMBAI -400056


April 2023

# ABSTRACT

Object detection methods are important in video surveillance applications to extract meaningful information from the images or video sequences. To identify objects in the video sequencing, Background subtraction methods are available but issues in the presence of Occlusion when object is obscured by other objects causing inaccurate localization of objects. The goal of this study is to provide a solution to of partial occlusion problem. The adaptive background subtraction method has been developed which aids in determining the initial threshold value for segmentation. The proposed background subtraction approach generates adaptive threshold values for each frame pixel-by-pixel for the accurate classification of foreground and background pixels. Due to shadow, few pixels are incorrectly categorized. To solve this problem, we proposed a intensity ratio based YCbcolor space model to determine the statistics of the pixels in the presence of shadow. The segmentation model was improved by using morphological operations and detecting the objects using bounding boxes which can be inaccurate localize in the presence of occlusion. We have proposed an improvement in axis aligned bounding box method to generate the enclosed bounding boxes around the objects for accurate localization even in the presence of occlusion. Next, we presented the occlusion prior condition to examine the statistics of pixels and determine whether or not they are occluded. Our approach is able to detect occlusions at various levels, ranging from 20% to 70%. Using benchmark datasets including PETS 2009 and Highway, we evaluated the performance of the proposed method in comparison to state-of-the-art approaches. Using IoU, GIoU, and CIoU indicators, the effectiveness of the suggested method has been assessed. Following testing, it was shown that the suggested approach can identify the object's visible components even when 65% of them are obscured. Findings from experimental work indicate that the strategy is effective in partial occlusion case detection and segmentation.

# CONTENTS

**ABSTRACT** ..... **i**

**LIST OF FIGURES** ..... **v**

**LIST OF TABLES** ..... **vii**

**CHAPTER 1** ..... **1**

INTRODUCTION ..... 1

1.1 Object Detection and Applications ..... 1

1.2 Background Subtraction and Challenges ..... 1

1.3 Research Motivation ..... 2

1.4 Problem Statement ..... 2

1.5 Research Objectives ..... 3

1.6 Research Contributions ..... 3

1.7 Organization of SYnopsis ..... 4

**CHAPTER 2** ..... **5**

LITERATURE REVIEW ..... 5

2.1 Background Subtraction ..... 5

2.2 Shadow Region Detection and Removal ..... 6

2.3 Occlusion in object Detection ..... 7

**CHAPTER 3** ..... **8**

ADAPTIVE BACKGROUND SUBTRACTION METHOD FOR MOVING OBJECT DETECTION ..... 8

3.1 Introduction ..... 8

3.2 Proposed Methodology ..... 8

3.3 Experimental results ..... 10

3.3.1 Qualitative Results ..... 10

3.3.2 Comparison of Qualitative Results of Proposed Method with State-of-the-art Methods....11
3.3.3 Quantitative Results....12
**CHAPTER 4....14**
INTENSITY RATIO BASED SHADOW REGION DETECTION AND REMOVAL APPROACH....14
4.1. Proposed Methodology....15
4.2. Experimental Setup....16
4.2.1. Datasets Description....16
4.2.2. Qualitative Results....16
4.2.3 Quantitative Results....17
**CHAPTER 5....19**
LOCALIZING AND DETECTING OBJECTS UNDER PARTIAL OCCLUSION....19
5.1 Proposed Methodology....19
5.1.1 View points for partial occlusion....21
5.1.2 Prior condition for partial occlusion detection....22
5.1.3 Various levels of occlusion....24
5.2 Experimental setup and results discussion....24
5.2.1 Datasets used....24
5.2.2 Occlusion measures....25
5.2.3 Qualitative results of proposed method....25
5.2.4 Quantitative results of proposed method....26
5.2.5 Bounding box localization loss for proposed method at different IoU threshold value....27
5.3 Comparative evaluation....29
5.3.1 Qualitative comparisons....29

5.3.2 Quantitative comparisons.....30

**CHAPTER 6.....32**

CONCLUSIONS AND FUTURE SCOPE.....32

**REFERENCES.....34**

**LIST OF PAPER PUBLICATIONS.....38**

# LIST OF FIGURES

**Fig. 1.1** Examples of object detection .....1
**Fig. 1.2** Process of Background Subtraction Method .....2
**Fig. 1.3** Proposed flow diagram of moving object detection under partial occlusion .....4
**Fig. 2.1** Few examples of Background Subtraction.....5
**Fig. 2.2** Basic process of Shadow Detection and Removal .....6
**Fig. 3.1** Basic Process of Background Subtraction Methods.....8
**Fig. 3.2** Flow diagram of proposed adaptive background subtraction approach.....9
**Fig. 3.3** The qualitative results of the proposed method on the 551st frame of the PETS2006 dataset .....10
**Fig. 3.4** The qualitative results of the proposed method on the 1592nd frame of the Highway dataset .....10
**Fig. 3.5** Qualitative comparative results of Highway dataset.....11
**Fig. 3.6** Qualitative comparative results of PETS2006 dataset .....11
**Fig. 4.1** Casting a shadow on the object .....14
**Fig. 4.2** General process of background subtraction and shadow removal approach.....14
**Fig. 4.3** Flow diagram of the proposed shadow region detection and removal approach .....15
**Fig. 4.4** The visual representation of the 1558th input frame on the Highway dataset .....17
**Fig. 4.5** Qualitative results of the 25$^{th}$ frame on SBU dataset under Test Case (I): Object is associated with shadow.....17
**Fig. 5.1** Different scenarios of occlusion occurrences.....19
**Fig. 5.2** Flow diagram of generating tighten box .....20
**Fig. 5.3 a & b** Visual representations before and after applying to tighten box method .....21
**Fig. 5.4** Localization of bounding boxes at different view point.....22
**Fig. 5.5** Flow diagram of detecting partial occlusion .....23
**Fig. 5.2** Flow diagram of generating tighten box .....20
**Fig. 5.6** Occlusion detection at Front-rear view on Highway dataset .....26
**Fig. 5.7** Occlusion detection at Left-right view on Highway dataset .....26
**Fig. 5.8** Occlusion detection at Top-down view on PETS2006 dataset .....26

**Fig. 5.9** Visual comparisons of proposed method with existing methods on the 209th frame of the Highway dataset.....29
**Fig. 5.10** Visual comparisons of proposed method with existing methods on the 429th frame of the PETS2006 dataset .....30

# LIST OF TABLES

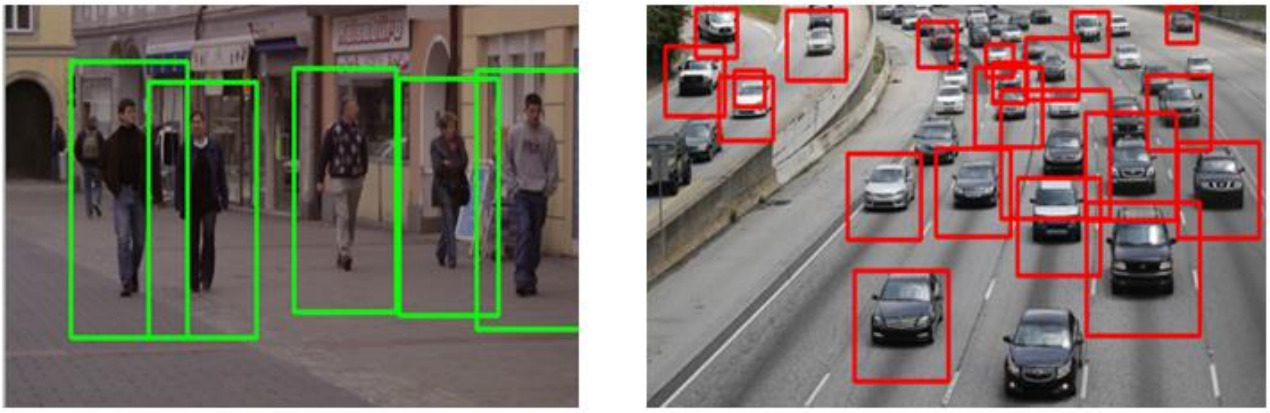
**Table 3.1** Description of Datasets .....10
**Table 3.2** Quantitative comparisons on PETS2006 dataset .....12
**Table 3.3** Quantitative comparisons on Highway dataset .....12
**Table 4.3** Quantitative results on Highway dataset for Test Case (I) .....18
**Table 4.4** Average comparative quantitative results on SBU dataset for Test Case (I) .....18
**Table 5.1** Different levels of occlusion [36] .....24
**Table 5.2** Ground truth frames and occluded test frames .....25
**Table 5.3** Ranges of occlusion measures .....25
**Table 5.4** Degrees of occlusion on Highway dataset .....27

# CHAPTER 1

# INTRODUCTION

## 1.1 OBJECT DETECTION AND APPLICATIONS

Object detection is a computer vision technique that aims to detect objects using machines, just like how humans do. It is an important research area to detect and localize the objects of interest within an image or video frame to extract meaningful information as shown in in Fig.1.1. The growth of object detection has led to increasing demand for real-time vision applications such as self driving, pedestrian detection, traffic analysis, and many more.



(a) Pedestrian detection (b) Car detection

**Fig. 1.1** Examples of object detection

The primary step of these real-time applications is the segmentation of interested regions by identifying the boundaries of objects within an image or video frame using Background Subtraction approaches for accurate detection of moving objects.

## 1.2 BACKGROUND SUBTRACTION AND CHALLENGES

In this method, foreground objects are considered moving objects, and the background is considered a static scene. The background subtraction process contains three stages as shown in Fig. 1.2 [1] (i) - Background initialization to generate the reference frame, (ii) – Foreground extraction using thresholding to segment the objects, and (iii) Background maintenance to update the background model.

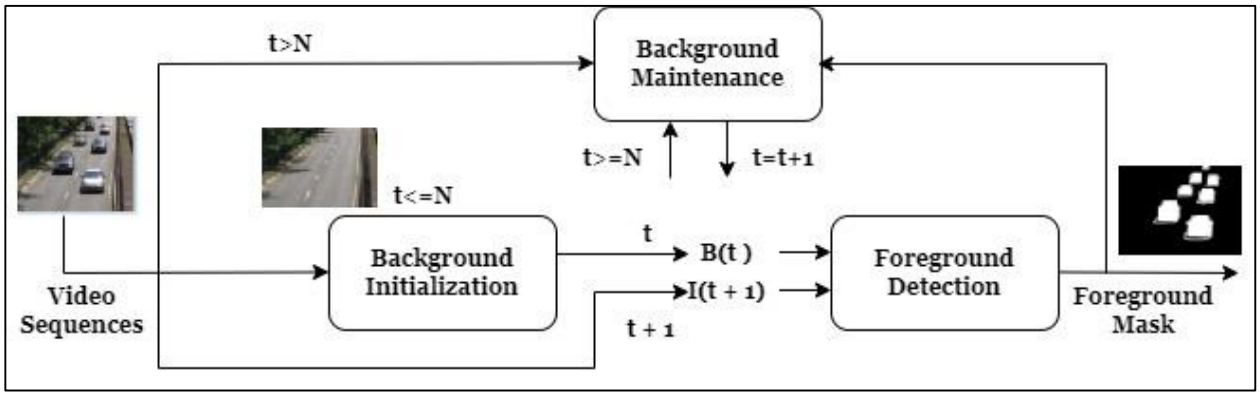



**Fig. 1.2** Process of Background Subtraction Method

Although the background subtraction process is effective in extracting foreground from video frames but facing the following challenges:

- Illumination changes
- Presence of shadow
- Clutter scene
- Camouflage
- Occlusion

## 1.3 RESEARCH MOTIVATION

Overcoming partial occlusion requires accounting for different viewpoints and background details. In our research, we focused on detecting moving objects in outdoor scenario with partial occlusion using a single, stable camera and a static background. We have started our research to keep in mind video surveillance applications, especially for vehicle and pedestrian detection. Hence, we used Highway and PETS2006 datasets [2] which contain outdoor video sequences captured in the daytime.

## 1.4 PROBLEM STATEMENT

Object intersection is the key point from where occlusion occurs. Detecting occluded objects and measuring the percentage of occlusion are the main challenges of object detection algorithms. To solve these challenges, the problem statement of this research work is:

*"To enhance the performance of detecting objects and measuring their overlap under partial occlusion conditions"*

## 1.5 RESEARCH OBJECTIVES

The following challenges have been observed for partial occlusion detection using background subtraction and shadow removal techniques: (i) choosing adaptive threshold value at the time of background model updation for foreground detection, (ii) misclassification of foreground/background pixels in the presence of object level shadows, (iii) Incorrect localization of occluded bounding boxes if the group in to one bounding box (iv) measuring the degree of overlapping in presence of partial occlusion.

To address these challenges, this thesis focuses on four Research Objectives (RO):

**$RO_1$ :** To detect and segment foreground objects using adaptive background subtraction method

**$RO_2$ :** To detect and remove shadow pixels in segmented results for improving the performance of proposed background subtraction method

**$RO_3$ :** To identify the correct localization of objects under the presence of partial occlusion

**$RO_4$ :** To measure the degree of overlap under partial occlusion condition

## 1.6 RESEARCH CONTRIBUTIONS

The flow diagram of the proposed work is shown by Fig.1.3 to detect objects and measure the overlapped degree under partial occlusion condition.

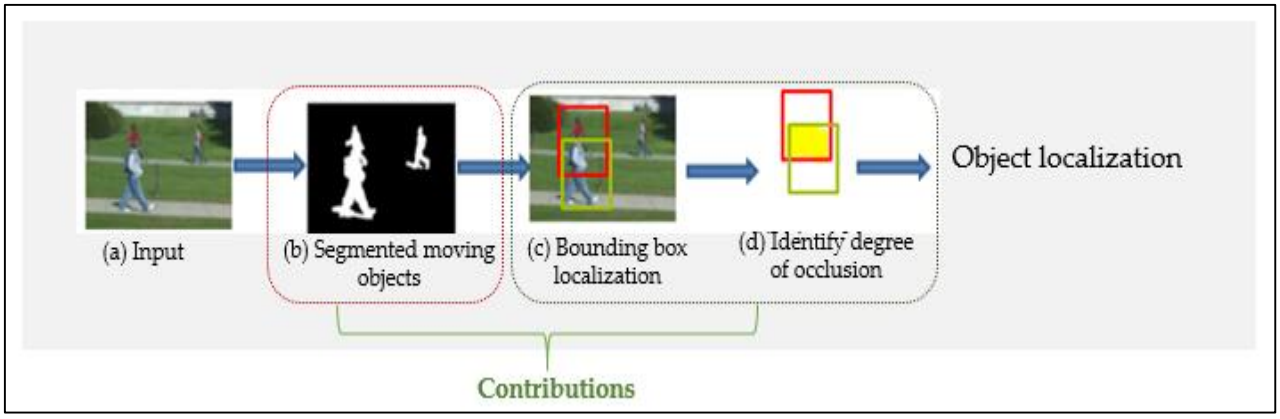



**Fig. 1.3** Proposed flow diagram of moving object detection under partial occlusion

The contributions of this research work are summarized below:

**$C_1$ :** Proposed a pixel-wise two-level adaptive background subtraction and updation method to generate the binary mask based on the automatic threshold value for detecting foreground objects.

**$C_2$ :** Proposed an intensity ratio-based YCbCr color space method for shadow region detection and removal from the segmented binary mask.

**$C_3$ :** Proposed an improved axis-aligned bounding box method using occlusion prior condition to localize the bounding boxes for identifying the objects under partial occlusion scenario.

**$C_4$ :** Evaluated and compared the occlusion measures metrics for measuring varying levels of occlusion at different viewpoints.

## 1.7 ORGANIZATION OF SYNOPSIS

- **Chapter 2:** Existing background subtraction methods and their limitations have been discussed. An overview of various shadow detection and removal methods has also been presented in this chapter. Then, different methods for occlusion detection have been deliberated.
- **Chapter 3:** Proposed background subtraction method has been discussed along with comparative analysis has also been presented in detail.
- **Chapter 4**: Proposed YCbCr color space-based shadow region detection and removal method is discussed along with different test cases. Then, the effectiveness of the proposed method is compared with state-of-the-art methods.
- **Chapter 5:** Proposed partial occlusion detection method is proposed to detect moving objects under occlusion conditions. Further, different levels of occlusion have been measured. Later, a comparative analysis of the proposed method is discussed.
- **Chapter 6**: Conclusions and future work of this research work have been discussed.

# CHAPTER 2

# LITERATURE REVIEW

This chapter highlights the key research findings and identifies gaps in the literature to conduct this research.

## 2.1 BACKGROUND SUBTRACTION

Background subtraction is the process of differentiating between foreground and background scenes using threshold values. It is a binary classification that classifies the pixels as either foreground (considering white pixels) or background (considering black pixels). The examples of background subtraction are shown in Fig. 2.1.

**Fig. 2.1** Few examples of Background Subtraction

To analyze the video sequences, moving object detection is the basic step to segment the foreground objects from the background. The most common method is Background subtraction methods have been proposed in the literature [3] [4] to segment foreground objects from the background using some threshold values. Some background subtraction methods generate misclassification of pixels due to the occurrence of shadow. The existence of shadows in segmented frames causes more false positives which may degrade the performance of the surveillance system. Many researchers have worked on detecting and removing shadow pixels using color space models to understand the color depth of the pixels [5][6]. These segmented results are helpful to understand the occlusion reasoning by correct localization of the detected objects. Few methods have been discussed in literature which achieve good results in a dynamic

environment, shadow removal, and illumination changes. Kernel Density Estimation, using GMM is proposed by [7] which discusses foreground identification and removal of shadow pixels using probability density functions (PDFs) for moving and stationary objects. Subsequently, a Codebook method is proposed with the Gaussian Mixture Model (GMM) to detect outliers using mean vectors and covariance matrices [8][9].

## 2.2 SHADOW REGION DETECTION AND REMOVAL

Background subtraction approaches are good to segment the objects but provides unsatisfactory results when shadows exist in the segment frame. The process of removing shadow pixels is shown in Fig. 2.3.

**Fig. 2.2** Basic process of Shadow Detection and Removal

Shadows deforms the geometric shape of moving objects, and may even cause the loss or merging of these objects[45][46]. Few methods are based on lightning conditions and illumination changes. The illumination-invariant based method reflect the object reflectivity[49] resulting in the loss of color information. A method developed in paper [51] to create a gradient field by calculating the mean and deviation of the gradients in these areas, they were able to recover a shadow-free image for the shadowed regions. Various methods have been proposed in the literature to identify shadow pixels and remove them using color correct methods[54][55]. A method proposed a hypothesis that shadows decrease the brightness and saturation of surfaces while preserving their hue in the HSV color space[54]. Methods have been proposed using the YUV color space to express variations in normalized luminance and chromaticity distortions [56][57]. The LAB color space method was proposed to detect and remove shadow region in dynamic background [58]. However, this method has limitations, such as the assumption that all luminance sources are white and that shadows and non-shadows have similar chrominance. To overcome these problems color constancy and sample based methods were proposed to deal with object level shadows but due to color variations false positives occurs [63] [11].

## 2.3 OCCLUSION IN OBJECT DETECTION

Various survey papers have been proposed in the literature to discuss existing methods to solve the occlusion problem [10][11]. When a few parts of the objects are obscured by other objects then this situation is known as partial occlusion. There are different types of occlusion, including inter-object occlusion and occlusion by the background scene structure. Self-occlusion can also occur when one part of an object obscures another, which is common in tracking articulated objects. However, in a scene with multiple objects, handling occlusion becomes more complicated. Multi-object tracking can be especially challenging when the targets are identical and described using the same model. The ability of tracking techniques to handle occlusion depends on the severity and state of the occlusion [12]. Few appearance based methods solve the occlusion issue generating pixel-wise masks for objects of interest, which accurately localize the objects and their shapes [13][14]. However, a method that uses hand-crafted features has been proposed to address this issue and detect curved objects even under occlusion. A similarity-based feature detection method was employed to detect occlusion by extracting histogram features using Local Binary Pattern. But fails to detect matching facial features. To address this limitation, a structure-aware key point tracking method was proposed using SVM based classifier to classify the vehicles[15] [16]. The other methods based on axis aligned bounding box methods[17] [18][19] are used to generate bounding box for determining the occlusion positions. The another method was proposed to handle occlusion based on adapted weighted strategy to detect the occluded objects [20]. It is important to measure the area of occluded objects and occlusion measure. Different similarity measures matrix were discussed in the literature. Intersection over Union is the first and foremost metrics to measure the overlapped objects[21][22].

Through this review, we have identified the research findings that remain to be addressed:

(i) Need of adaptive thresholding for background subtraction method

(ii) Automatic removal of shadow pixels in segmented frame using color feature and

(iii) Partial occlusion detection by measuring degree of occlusion.

The next chapters explained the proposed methodologies to solve the addresses challenges.

# CHAPTER 3

# ADAPTIVE BACKGROUND SUBTRACTION METHOD FOR MOVING OBJECT DETECTION

## 3.1 INTRODUCTION

The Background Subtraction method is the most commonly used approach for moving object detection. It computes the difference between the current frame and a reference frame (without moving object) on a pixel-by-pixel basis and then chooses the threshold value (T) to generate a binary mask that distinguishes between moving foreground objects and stationary background objects as depicted in Fig. 3.1 [23].

**Fig. 3.1** Basic Process of Background Subtraction Methods

## 3.2 PROPOSED METHODOLOGY

A two-level pixel-wise adaptive thresholding method has been proposed to solve the problem of choosing threshold values for foreground and background segmentation in static environments using a single camera. Also, the update mechanism is used to segment foreground objects in new frames by comparing the pixel values of each frame to the background model. This stage helps in recognizing the foreground objects when they appear in the frame. Further, the proposed process can be modified by using post-processing operations such as morphological operations to improve the accuracy of the segmentation. The flow diagram of the proposed work is represented in Fig. 3.2.

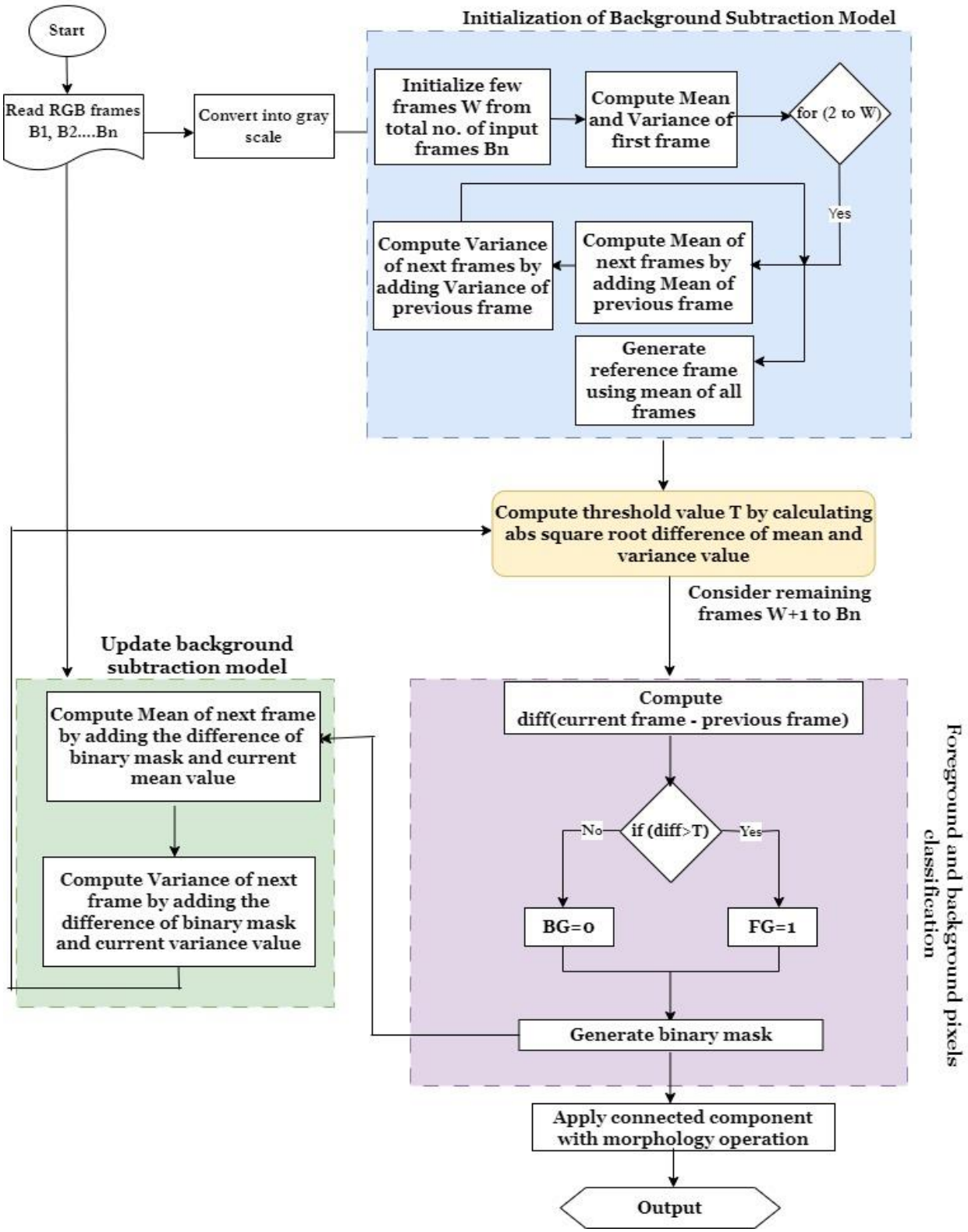

Start
Read RGB frames B1, B2....Bn
Convert into gray scale
Initialization of Background Subtraction Model
Initialize few frames W from total no. of input frames Bn
Compute Mean and Variance of first frame
for (2 to W)
Yes
Compute Variance of next frames by adding Variance of previous frame
Compute Mean of next frames by adding Mean of previous frame
Generate reference frame using mean of all frames
Compute threshold value T by calculating abs square root difference of mean and variance value
Consider remaining frames W+1 to Bn
Update background subtraction model
Compute Mean of next frame by adding the difference of binary mask and current mean value
Compute Variance of next frame by adding the difference of binary mask and current variance value
Compute diff(current frame - previous frame)
No
if (diff>T)
Yes
BG=0
FG=1
Generate binary mask
Foreground and background pixels classification
Apply connected component with morphology operation
Output

**Fig. 3.2** Flow diagram of proposed adaptive background subtraction approach

## 3.3 EXPERIMENTAL RESULTS

The proposed Background Subtraction method has been evaluated on two widely used datasets, namely Highway [2] and PETS 2006 [2] as described in Table 3.1. The experimental findings are presented in this section, demonstrating the effectiveness of the proposed method through qualitative (visual perception) and quantitative analysis based on available ground truth frames.

**Table 3.1** Description of Datasets

| **Datasets** | **Total number frames** | **Size of frame** | **Ground Truth frames** | **Objects** |
|---|---|---|---|---|
| PETS2006 | 1200 | 720 x 576 | 1200 | Pedestrian |
| Highway | 1700 | 320 x 240 | 200 | Vehicle |

### 3.3.1 Qualitative Results

The qualitative results of the proposed method are shown in Fig. 3.3 for the PETS2006 dataset and in Fig. 3.4 for the Highway dataset. It has been observed the grayscale range of 40 to 50 pixels through experimental analysis to produce a reference frame. This threshold value is used for pixel classification then morphology operations are applied to generate the bounding box on each detected object.

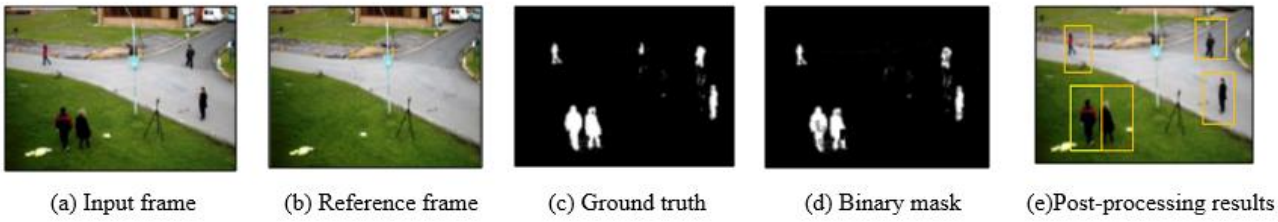



**Fig. 3.3** The qualitative results of the proposed method on the 551st frame of the PETS2006 dataset

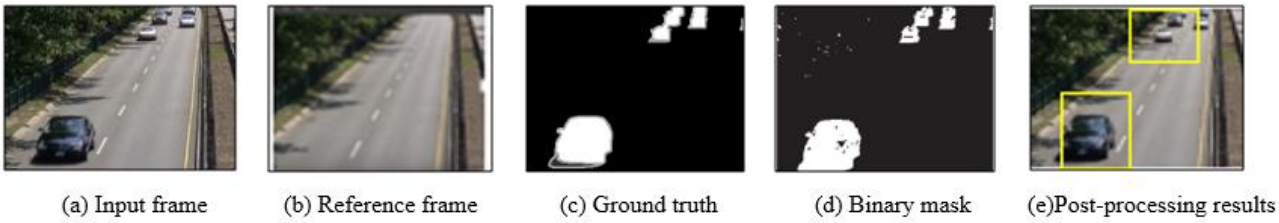



**Fig. 3.4** The qualitative results of the proposed method on the 1592nd frame of the Highway dataset

### 3.3.2 Comparison of Qualitative Results of Proposed Method with State-of-the-art Methods

The qualitative results of the proposed method have been compared with five state of the art methods such as (i) Temporal Average Filter (TAF) [2], (ii) Running Average Gaussian (RAG) [24], (iii) Adaptive Gaussian Mixture Model (AGMM)[8], (iv) Weighted Patch-Based Background Subtraction (WePBAS) [10], and (v) Visual Background Extractor (ViBe) [9]. The comparative qualitative results have been shown in Fig. 3.5 and Fig. 3.6 respectively.

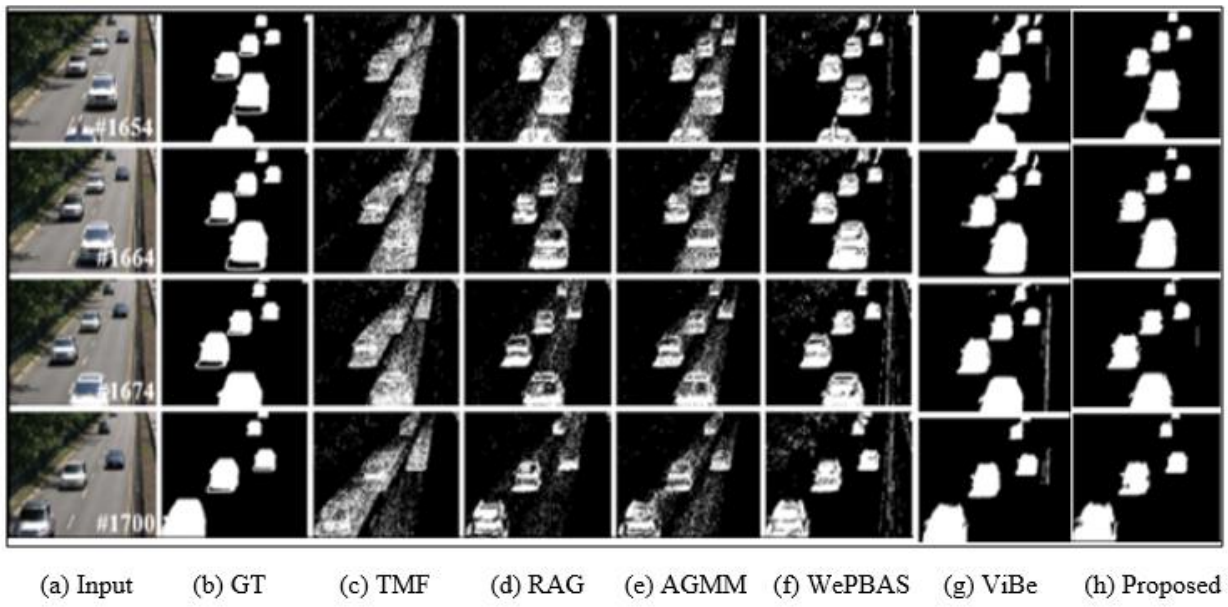



**Fig. 3.5** Qualitative comparative results of Highway dataset

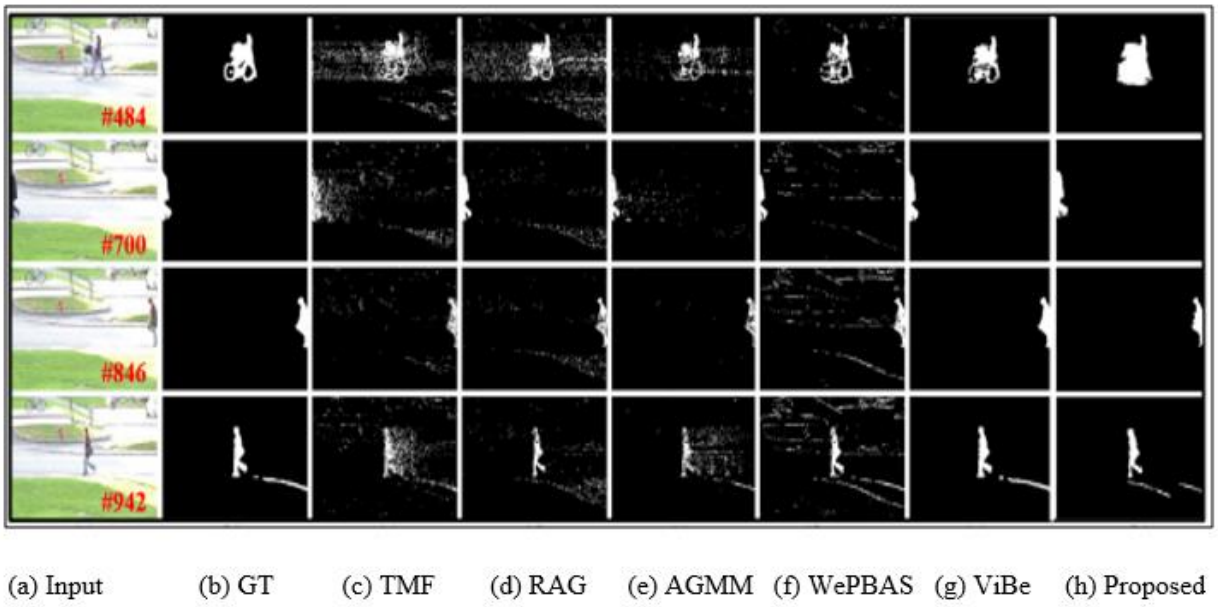



**Fig. 3.6** Qualitative comparative results of PETS2006 dataset

### 3.3.3 Quantitative Results

The proposed method is evaluated using six statistical measures, such as Precision, Recall, FNR, FPR, F-Measure, and PCC, all of which are sensitive to the reference frame. The qualitative results are shown in Table 3.2 and 3.3 respectively.

**Table 3.2** Quantitative comparisons on PETS2006 dataset

| **Methods** | **Precision ↑** | **Recall ↑** | **FPR ↓** | **FNR ↓** | **F-Measure ↑** | **PCC ↑** |
|---|---|---|---|---|---|---|
| TMF[4] | 0.5127 | 0.5284 | 0.3261 | 0.1002 | 0.5149 | 0.6872 |
| RAG [24] | 0.5893 | 0.6255 | 0.3171 | 0.1142 | 0.5993 | 0.7436 |
| AGMM [25] | 0.8303 | 0.7282 | 0.1615 | 0.0662 | 0.7934 | 0.8364 |
| WePBAS[26] | 0.8481 | 0.7615 | 0.1063 | 0.0216 | 0.8195 | 0.9053 |
| ViBe[27] | 0.8790 | 0.8191 | 0.0657 | 0.0153 | 0.8654 | 0.9342 |
| Proposed | **0.9496** | **0.8716** | **0.0341** | **0.0068** | **0.9449** | **0.9646** |

Bold values illustrate better results.

**Table 3.3** Quantitative comparisons on Highway dataset

| **Methods** | **Precision ↑** | **Recall ↑** | **FPR ↓** | **FNR ↓** | **F-Measure ↑** | **PCC ↑** |
|---|---|---|---|---|---|---|
| TMF[4] | 0.5024 | 0.5358 | 0.3083 | 0.1045 | 0.5145 | 0.6683 |
| RAG [24] | 0.5762 | 0.6057 | 0.3054 | 0.1501 | 0.5846 | 0.7162 |
| AGMM [25] | 0.8182 | 0.6948 | 0.2957 | 0.1654 | 0.7452 | 0.7296 |
| WePBAS[26] | 0.8250 | 0.7661 | 0.1092 | 0.1203 | 0.8047 | 0.8992 |
| ViBe[27] | 0.8590 | 0.7525 | 0.0765 | 0.0788 | 0.8542 | 0.9221 |
| Proposed | **0.9254** | **0.9041** | **0.0434** | **0.0743** | **0.9247** | **0.9573** |

Experimental findings indicate that the range of initial threshold values is between 50 and 70 pixels for highway dataset and the adaptive threshold value range between 90 and 125 pixels. The effectiveness of the proposed method is evaluated using six statistical parameters on two benchmark datasets, Highway and PETS2006. Results demonstrate that the proposed model outperforms existing methods in terms of both quantity and quality. Chapter 3 of this thesis discusses our second contribution, which focuses on the shadow region detection and removal process.

# CHAPTER 4

## INTENSITY RATIO BASED SHADOW REGION DETECTION AND REMOVAL APPROACH

Background subtraction techniques are efficient in identifying moving objects; however, they tend to incorrectly classify pixels when shadows are present. The existence of shadow distorts the shape of the detected object causing wrong recognition of the object as shown in Fig. 4.1. The general process of background subtraction and shadow removal is shown in Fig. 4.2.

(a) Object with shadow (b) Binary mask

**Fig. 4.1** Casting a shadow on the object

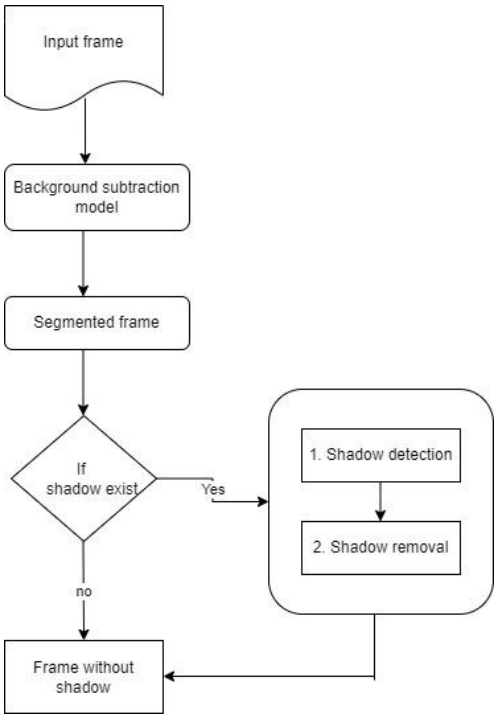



**Fig. 4.2** General process of background subtraction and shadow removal approach

## 4.1. PROPOSED METHODOLOGY

We have proposed an intensity ratio-based YCbCr color space method for shadow region detection and its removal. To identify the presence of shadows, our model uses a binary mask generated by the background subtraction model and a color frame of the foreground. We have used YCbCr based color correction method to correct the luminance (brightness) values generated by the Y color channel and chrominance values generated by Cb and Cr color channels. The flow graph of the proposed method is shown in Fig. 4.3.

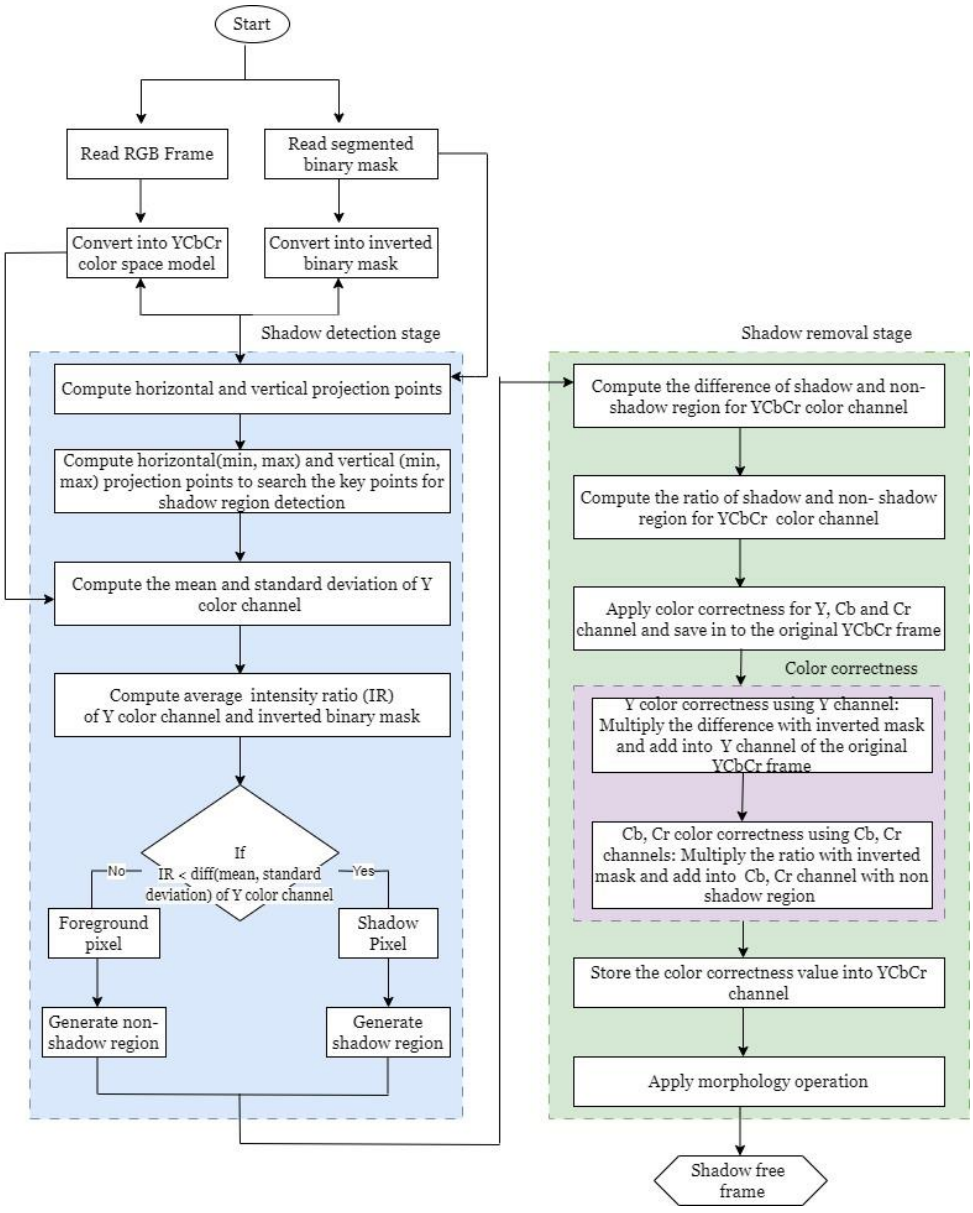

Start
Read RGB Frame
Read segmented binary mask
Convert into YCbCr color space model
Convert into inverted binary mask
Shadow detection stage
Compute horizontal and vertical projection points
Compute horizontal(min, max) and vertical (min, max) projection points to search the key points for shadow region detection
Compute the mean and standard deviation of Y color channel
Compute average intensity ratio (IR) of Y color channel and inverted binary mask
If IR < diff(mean, standard deviation) of Y color channel
No
Yes
Foreground pixel
Shadow Pixel
Generate non-shadow region
Generate shadow region
Shadow removal stage
Compute the difference of shadow and non-shadow region for YCbCr color channel
Compute the ratio of shadow and non- shadow region for YCbCr color channel
Apply color correctness for Y, Cb and Cr channel and save in to the original YCbCr frame
Color correctness
Y color correctness using Y channel: Multiply the difference with inverted mask and add into Y channel of the original YCbCr frame
Cb, Cr color correctness using Cb, Cr channels: Multiply the ratio with inverted mask and add into Cb, Cr channel with non shadow region
Store the color correctness value into YCbCr channel
Apply morphology operation
Shadow free frame

**Fig. 4.3** Flow diagram of the proposed shadow region detection and removal approach

## 4.2. EXPERIMENTAL SETUP

We considered three test cases (as per the dataset) to evaluate the performance of the proposed method:

- Test Case 1 – Object is associated with shadow
- Test Case 2 – Only shadow is available
- Test Case 3 – Multiple shadows are available

### 4.2.1. Datasets Description

We have used two benchmark datasets (i) Highway[2] and (ii) SBU [28] to test the effectiveness of the proposed method. The Highway dataset has also been used for evaluating the performance of the proposed background subtraction method as discussed in Chapter 2. The SBU dataset is specially designed for shadow detection and removal in which every image is associated with shadow. Table 4.1 shows test cases for particular datasets.

**Table 4.1** Applicable Test Cases and Associated Frames

| Datasets | Test cases | Sample Frames |
|---|---|---|
| Highway | Test Case (I) - few frames are associated with shadow | Total 45 frames are associated with shadow |
| SBU | Test Case (I), Test Case (II), Test Case (III) (Shadow based dataset) | Test Case (I) – 200 frames<br>Test Case (II) – 250 frames<br>Test Case (III) – 100 frames |

### 4.2.2. Qualitative Results

Fig. 4.4 demonstrates the results of the $1558^{th}$ input frame from the Highway dataset. The keypoints for shadow pixels are generated and highlighted those keypoints in yellow color as shown in 4(c).After that shadow free image is generated as shown in Fig.4.4 (e). The qualitative results for Test Case (II) and Test Case (III) were also evaluated (presented in thesis). While removing the shadow pixels, a few internal holes exist which have been removed using post-processing operations.

(a) Input frame (b) Generated binary mask (c) Key points of shadow pixels (d) Detection of shadow region (e) Shadow-free image

**Fig. 4.4** The visual representation of the 1558th input frame on the Highway dataset

Similarly, the proposed method has been applied on the SBU dataset. The qualitative results for the same have been shown in Fig. 4.5, Fig. 4.6 and Fig. 4.7 for different Test Cases (I), (II) and (III) respectively.

(a) Input frame (b) Generated binary mask (c) Key points of shadow pixels (d) Detection of shadow region (e) Shadow-free image

**Fig. 4.5** Qualitative results of the 25$^{th}$ frame on SBU dataset under Test Case (I): Object is associated with shadow

### 4.2.3 Quantitative Results

To evaluate the effectiveness of the proposed method, two performance measures are used: (i) the shadow detection rate ($\eta$), which indicates the percentage of shadow pixels correctly identified, and (ii) the shadow discrimination rate ($\xi$), which measures the ability to distinguish between foreground objects and shadow regions by removing the shadow pixels. The evaluation included three test cases: (I) when the shadow is associated with objects, (II) when only the shadow is present, and (III) when multiple shadows are present. The PETS 2006 dataset was excluded from the evaluation as it did not contain shadow pixels. Only Test (I) results are displayed here. In the Highway dataset, object visibility is important and shadow is associated with an object. The average comparative quantitative results of the proposed method with the existing method have been shown in Table 4.3 of the Highway dataset which is applicable for Test Case (I) for shadow removal.

**Table 4.1** Quantitative results on Highway dataset for Test Case (I)

| **Methods** | **Shadow detection rate (η %)** | **Shadow removal rate (ξ %)** |
|---|---|---|
| LAB Method[29] | 82.17% | 74.11% |
| HSV Method[30] | 82.02% | 72.52% |
| Sample Based Method[31] | 86.26% | 83.23% |
| Color Constancy Method[32] | 87.41% | 91.21% |
| Proposed Method | **92.57%** | **93.16%** |

The bold value represents good results.

The SBU dataset is specifically designed to detect and remove shadows. Table 4.4, Table 4.5, and Table 4.6 illustrate the average quantitative evaluation of the proposed method when compared to other methods for Test Case (I).

**Table 4.2** Average comparative quantitative results on SBU dataset for Test Case (I)

| **Methods** | **Shadow detection rate (η %)** | **Shadow discrimination rate (ξ %)** |
|---|---|---|
| LAB Method[29] | 79.93% | 75.68% |
| HSV Method[30] | 78.86% | 72.20% |
| Sample Based Method[31] | 86.82% | 84.43% |
| Color Constancy Method[32] | 90.03% | 89.11% |
| Proposed Method | **94.04%** | **95.81%** |

Experiments revealed that the shadow region pixels range lie from 40 to 80 pixels values. Later morphology and connected component operations were applied to detect the objects with bounding boxes. The next chapter discusses

# CHAPTER 5

# LOCALIZING AND DETECTING OBJECTS UNDER PARTIAL OCCLUSION

In the real world, multiple objects exist in a particular scene and it is obvious that the visibility of one object is hidden due to another object. This situation is called occlusion as shown in Fig. 5.1. One of the crucial challenges of object detection algorithms is to accurate detection of multiple objects under an occluded environment. There can be either partial or full occlusion exists.

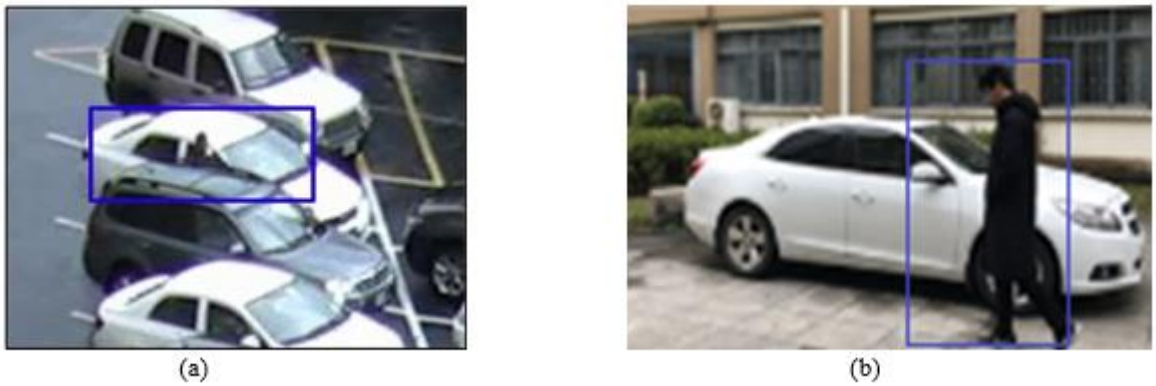



**Fig. 5.1** Different scenarios of occlusion occurrences [33]

As per the literature the following research gaps were found:

- Under partial occlusion conditions, if objects are in close proximity to each other, generating a single bounding box instead of multiple bounding boxes can lead to incorrect localization
- Lack of standardized methods to measure the overlapped area of partially detected objects at different levels of occlusion

The proposed method solve the above challenges by proposing Occluded Bounding Box Detector based on an axis-aligned bounding box approach with occlusion prior condition to automatically detect partially occluded objects in recorded scenes. We also explored the effectiveness of the Intersection over Union (IoU) similarity measure metric to detect overlapping between two objects[34][35].

## 5.1 PROPOSED METHODOLOGY

This section presents a detailed discussion of the methodology employed in the proposed research. Firstly, an enclosed boundary box is generated on the input sequence to detect the region of interest called objects

using axis-aligned bounding box (AABB) method to generate the enclosed boundary box on the detected objects. The next step is to check the overlapped bounding boxes for occlusion detection. If there are multiple boundary boxes that are overlapped they can be identified as occluded objects by using the improved Intersection Over Union (IOU) metric otherwise detected objects are considered non-occluded objects. We also calculate the percentage % of the overlapping area of bounding boxes to understand the visible features of occluded objects. The general flow of the proposed work is illustrated in Fig. 5.2.

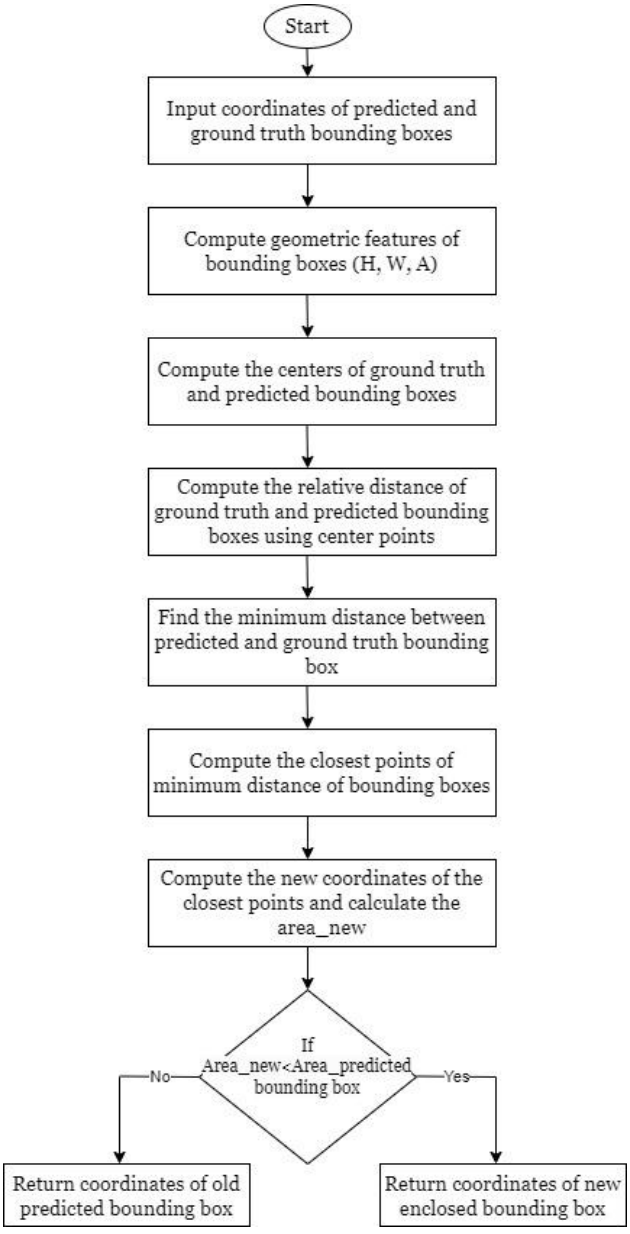
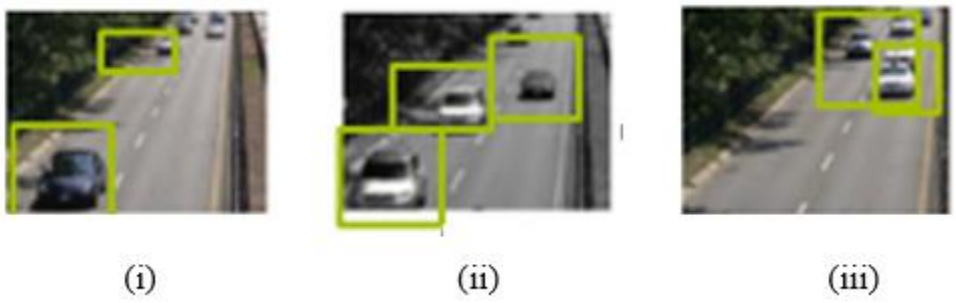
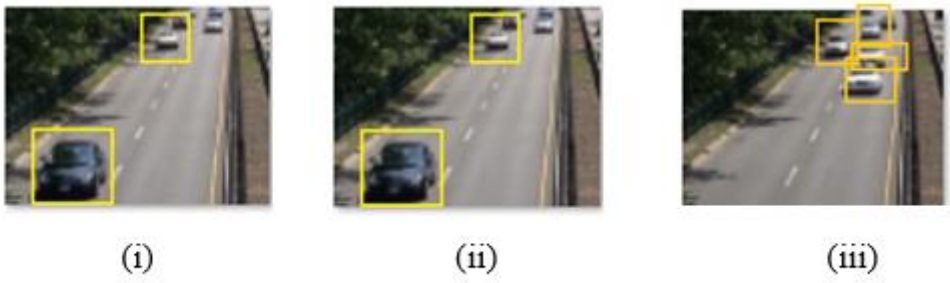



**Fig. 5.2** Flow diagram of generating tighten bounding box

This method helps in generating the bounding box for those objects which are partially occluded and one bounding box was generated before applying tighten box method. Fig. 5.3 (a) and Fig. 5.3 (b) illustrate the outcomes generated in which tighten bounding boxes are displayed on moving objects such as cars.




**(a)** Visual representations before applying to tighten box method




**(b)** Visual representations after applying to tighten box method

**Fig. 5.3 (a) & (b)** Visual representations before and after applying to tighten box method

### 5.1.1 View points for partial occlusion

We have made a few assumptions before applying the occlusion detection method, such that the camera is static and has a specific viewpoint. Our research proposes a method for occlusion detection using a static camera with multiple viewpoints. It is essential to note that only one viewpoint can be considered at a time. In this research work, we have used three different viewpoints: (i) left-right, (ii) top-down, and (iii) front-rear, along with localization of bounding boxes in the occurrence of partial occlusion scenario as shown in Fig. 5.4..

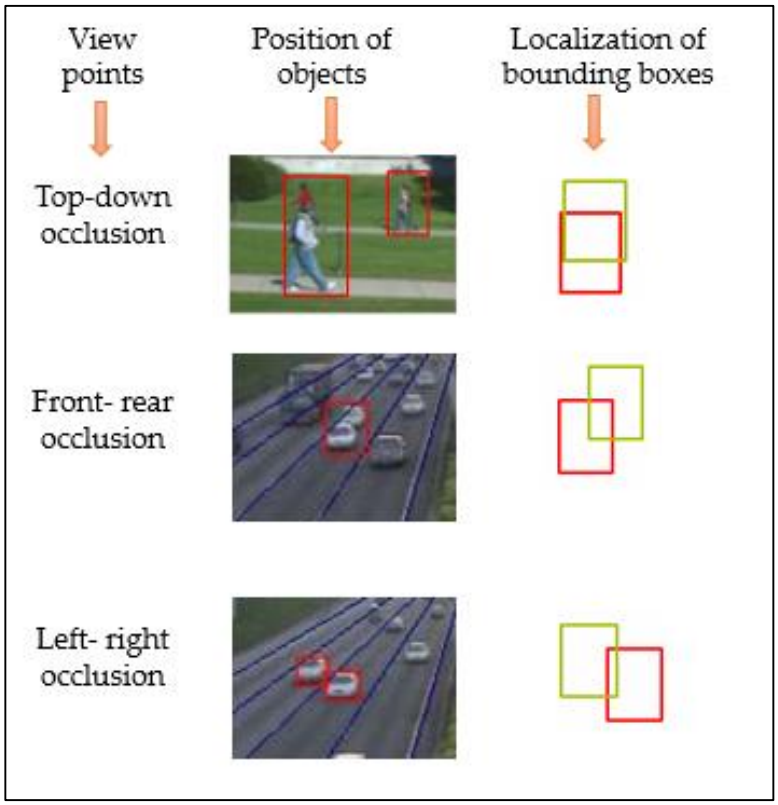



**Fig. 5.4** Localization of bounding boxes at different view point

### 5.1.2 Prior condition for partial occlusion detection

Object orientation is also important for occlusion detection on moving objects to identify which object is occluded to other objects. Once the coordinates of the enclosed bounding box have been generated, the step is to determine whether there is an occluded region. The process of detecting partial occlusion is represented by Fig. 5.5.

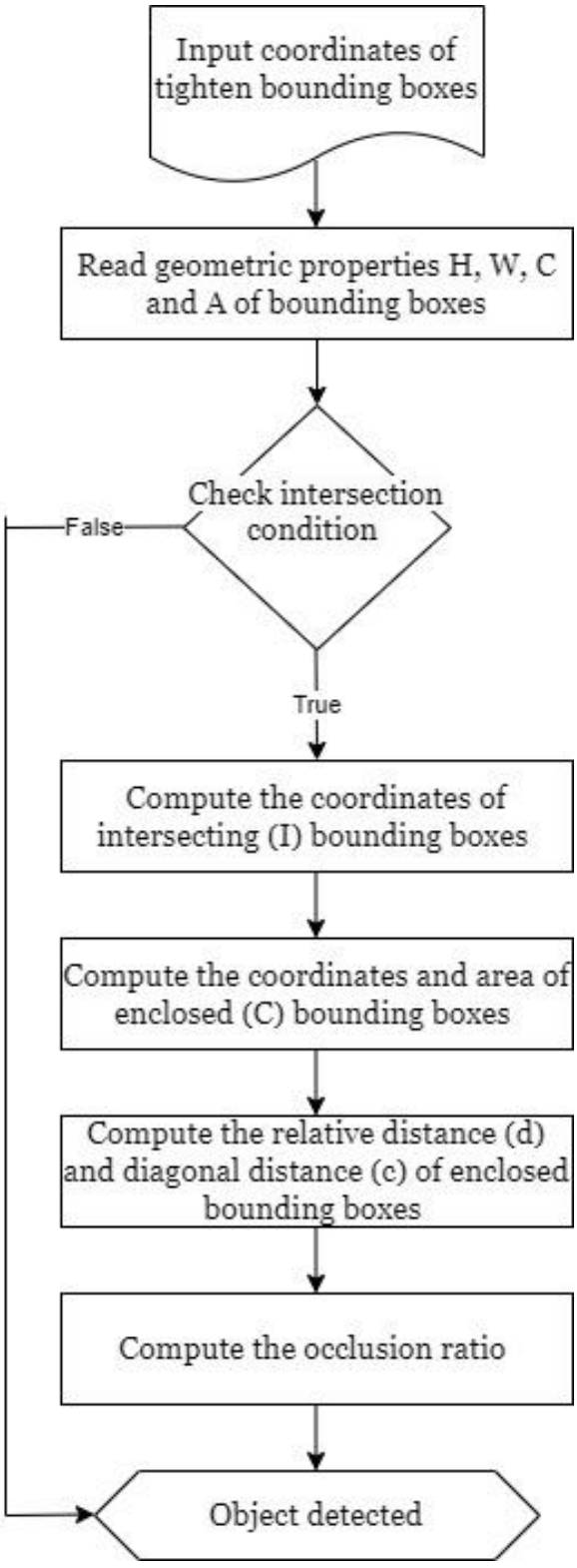



**Fig. 5.5** Flow diagram of detecting partial occlusion

### 5.1.3 Various levels of occlusion

Levels of occlusion can determine how much an object is obstructed by other objects and the impact of occlusion level on object visibility. Table 5.1 describes varying levels of occlusion along with the visibility part of the object. We have used some ranges of an overlapping area such as 20% - 40% and 40% - 70% based on detected bounding boxes for partial occlusion.

**Table 5.1** Different levels of occlusion [36]

| **Levels of occlusion** | **Overlapping range** | **Visibility criteria of object** |
|---|---|---|
| No occlusion | 0% - 10% | More features of objects are visible |
| Low-level occlusion | 10% - 20% | More features of objects are visible |
| Partial occlusion | 20 % - 40% & 40% - 70% | Visibility of object is less causing false positives and bounding box localization error |
| Heavy occlusion | >70% | The visibility of an object is very less |

## 5.2 EXPERIMENTAL SETUP AND RESULTS DISCUSSION

In this section, we discuss the findings of our experiments which aimed to evaluate the effectiveness of the proposed method for detecting objects that are partially overlapped within video sequences.

### 5.2.1 Datasets used

To detect overlapping regions during partial occlusion, we used two benchmark datasets: Highway and PETS 2006[2]. For experimentation purposes, we considered a few frames from both datasets under different levels of occlusion, as illustrated in Table 5.2.

***Table 5.2*** Ground truth frames and occluded test frames

| **Datasets** | **Ground truth frames** | **View points** | **Total number of occluded frames in different levels of occlusion** | |
|---|---|---|---|---|
| | | | **20% - 40%** | **40% -70%** |
| Highway | 200 frames | Front rear view & Left right view | 45 frames | 52 frames |
| PETS2006 | 1200 frames | Top down view | 400 frames | 456 frames |

### 5.2.2 Occlusion measures

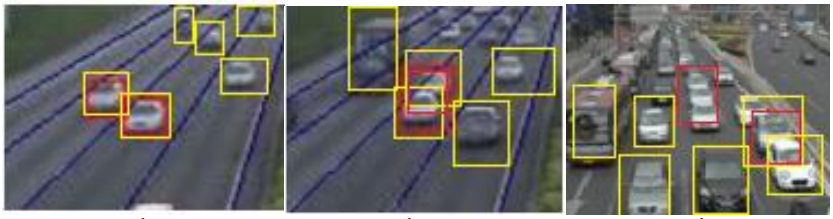
Metrics for evaluating occlusion are used to measure the efficiency of methods intended to detect and handle occlusions in video sequences. Table 5.3 represents different occlusion measures for identifying occlusion and non-occlusion scenarios where 0 represents the nonexistence of occlusion while 1 represents full occlusion.

**Table 5.3** Ranges of occlusion measures

| **Measures of occlusion** | **Ranges** | **Non-Occlusion/ Occlusion** |
|---|---|---|
| Intersection over Union (IoU) | 0 to 1 | 0: Non occlusion, 1: full occlusion |
| Generalized Intersection over Unio (GIoU) | -1 to 1 | -1: Non occlusion, 1: full occlusion |
| Complete Intersection over Union (CIoU) | - 1 to 1 | -1: Non occlusion, 1: full occlusion |

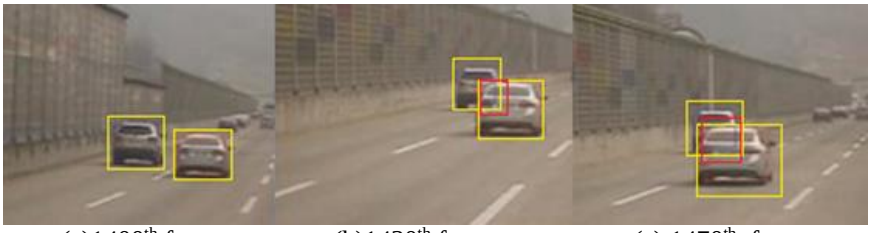
### 5.2.3 Qualitative results of proposed method

The visual representations of the proposed method have been shown in Fig. 5.6, Fig. 5.7, and Fig. 5.8 for Highway and PETS2006 datasets. The identified objects are outlined with a yellow bounding box, while the occluded regions are indicated with a red bounding box.



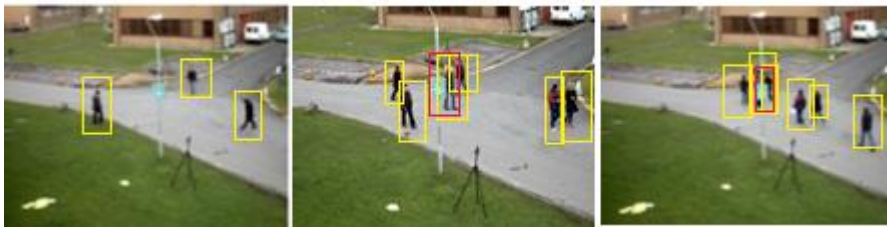
(a)1502th frame <15% occlusion | (b)1526th frame 30% occlusion | (c)1550th frame 56% , 24% occlusion

**Fig. 5.6** Occlusion detection at Front-rear view on Highway dataset



(a)1400th frame 0% occlusion | (b)1430th frame 23% occlusion | (c) 1470th frame 50% occlusion

**Fig. 5.7** Occlusion detection at Left-right view on Highway dataset



(a) 120th frame 0% occlusion | (b) 155th frame 22% occlusion | (c) 190th frame 65% occlusion

**Fig. 5.8** Occlusion detection at Top-down view on PETS2006 dataset

### 5.2.4 Quantitative results of proposed method

In this section, we discuss quantitative analysis of the proposed method using three occlusion evaluation metrics, namely Intersection over Union (IoU), Generalized Intersection over Union (GIoU), and Complete Intersection over Union (CIoU). These metrics are used to measure the statistics of pixel value under partial occlusion. The quantitative analysis for the proposed method for same frames is depicted in Table 5.4 and 5.5.

**Table 5.4** Degrees of occlusion on Highway dataset

| View point | Sample frames | Non-Occlusion | | | Occlusion | | |
|---|---|---|---|---|---|---|---|
| | | **IoU** | **GIoU** | **CIoU** | **IoU** | **GIoU** | **CIoU** |
| **Highway Dataset** | | | | | | | |
| View Point *Left-right* | 1400th frame | 0 | -1 | -1 | --- | --- | --- |
| | 1430th frame | --- | --- | --- | 0.230 | 0.229 | 0.227 |
| | 1470th frame | --- | --- | --- | 0.50 | 0.493 | 0.487 |
| View Point *Front-rear* | 1502th frame | --- | --- | --- | 0.147 | 0.142 | 0.140 |
| | 1526th frame | --- | --- | --- | 0.30 | 0.283 | 0.275 |
| | 1550th frame | --- | --- | --- | 0.561, 0.23 | 0.559, 0.228 | 0.548, 0.221 |
| **PETS2006 Dataset** | | | | | | | |
| Top Down | 120th frame | 0 | -1 | -1 | --- | --- | --- |
| | 155th frame | 0.051 | -01 | -01 | 0.220 | 0.219 | 0.216 |
| | 190th frame | --- | --- | --- | 0.657 | 0.624 | 0.621 |

According to the experimental results, the use of Intersection over Union (IoU) metric for occlusion detection represent occlusion ratio. Based on IoU, GIoU and CIoU values have been measured and it should be close to IoU.

### 5.2.5 Bounding box localization loss for proposed method at different IoU threshold value

To assess the accuracy of object detection systems, Bounding Box Localization (BBL) loss metrics are used, which compare predicted and ground truth bounding boxes to identify true and false positives. Loss represents misalignment of bounding boxes. It is based on IoU threshold value ($\delta$)

Table 5.5 represent that our proposed method achieved a relative loss improvement of 14.2% in GIoU loss and 17.6% in CIoU loss under 65% occlusion. Table 5.6 displays the results of the proposed

method on different threshold values for occlusion detection. The findings reveal that when there is 65% occlusion, the proposed method improves the CIoU loss by an average of 17.3%.

**Table 5.5** Quantitative comparison of BBL loss at different threshold values (δ) on Highway dataset

| **BBL Loss** | **IoU Threshold (δ)** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***0.25*** | ***0.30*** | ***0.35*** | ***0.40*** | ***0.45*** | ***0.50*** | ***0.55*** | ***0.60*** | ***0.65*** | ***0.70*** | $Avg_{Loss}$ |
| IoU | 0.231 | 0.286 | 0.342 | 0.385 | 0.438 | 0.492 | 0.537 | 0.583 | 0.629 | 0.673 | 0.45 |
| $L_{IoU}$ | 0.019 | 0.014 | 0.008 | 0.015 | 0.012 | 0.008 | 0.013 | 0.017 | 0.021 | 0.027 | 0.015 |
| $L_{GIoU}$ | 0.015 | 0.013 | 0.006 | 0.012 | 0.009 | 0.007 | 0.010 | 0.014 | 0.018 | 0.022 | **0.012** |
| RI% | 0.163 | 0.071 | 0.250 | 0.20 | 0.250 | 0.050 | 0.230 | 0.176 | 0.142 | 0.185 | **0.171** |
| $L_{CIoU}$ | 0.016 | 0.012 | 0.010 | 0.011 | 0.010 | 0.006 | 0.009 | 0.012 | 0.017 | 0.021 | 0.012 |
| RI% | 0.157 | 0.135 | -0.25 | 0.266 | 0.166 | 0.187 | 0.253 | 0.247 | 0.176 | 0.222 | 0.156 |

RI = Relative Improvement, δ = overlapped threshold value representing degree of occlusion. Bold values depict better results. Negative value represents a misalignment of bounding box

**Table 5.6** Quantitative comparison of BBL loss on different overlapped threshold values (δ) on PETS 2006 dataset

| ***BBL Loss*** | ***IoU Threshold (δ)*** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **0.25** | **0.30** | **0.35** | **0.40** | **0.45** | **0.50** | **0.55** | **0.60** | **0.65** | **0.70** | $Avg_{Loss}$ |
| $IoU$ | 0.229 | 0.277 | 0.327 | 0.375 | 0.427 | 0.479 | 0.531 | 0.588 | 0.627 | 0.674 | 0.453 |
| $L_{IoU}$ | 0.021 | 0.023 | 0.023 | 0.025 | 0.023 | 0.021 | 0.019 | 0.012 | 0.023 | 0.026 | 0.021 |
| $L_{GIoU}$ | 0.019 | 0.020 | 0.018 | 0.024 | 0.019 | 0.017 | 0.018 | 0.009 | 0.020 | 0.023 | 0.018 |
| $RI\%$ | 0.095 | 0.100 | 0.217 | 0.032 | 0.173 | 0.161 | 0.021 | 0.250 | 0.130 | 0.088 | 0.124 |
| $L_{CIoU}$ | 0.018 | 0.020 | 0.017 | 0.021 | 0.017 | 0.016 | 0.016 | 0.010 | 0.019 | 0.022 | **0.017** |
| $RI\%$ | 0.142 | 0.130 | 0.239 | 0.160 | 0.243 | 0.219 | 0.131 | 0.167 | 0.173 | 0.153 | **0.176** |

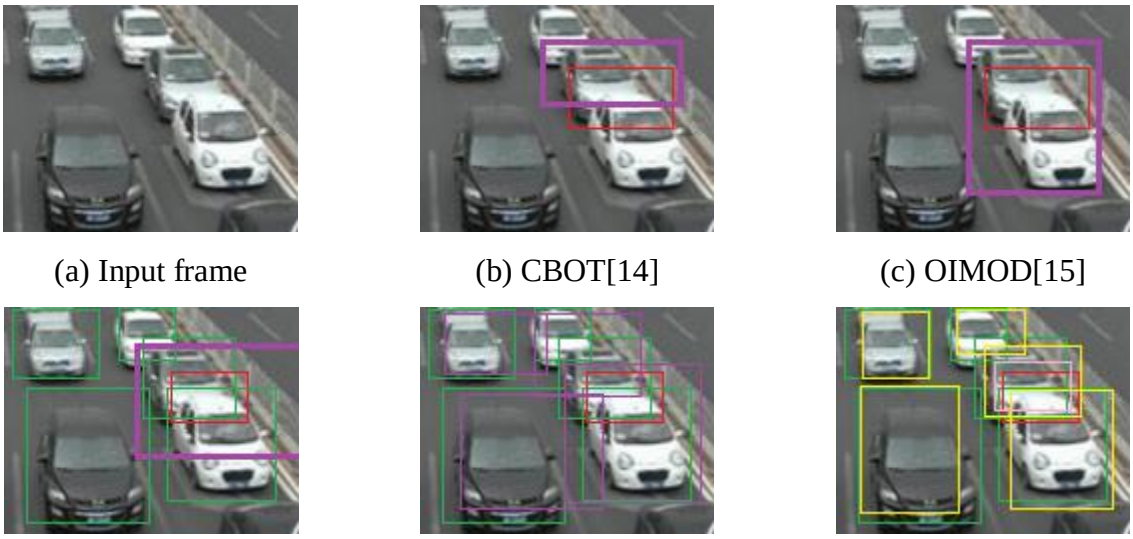
RI = Relative Improvement. Bold values depict better results.

## 5.3 COMPARATIVE EVALUATION

The results have been compared with four state-of-the-art methods (i) CBOT[14], (ii) OIMOD[15], (iii) SAOD[37], and (iv) AWOD[18].

### 5.3.1 Qualitative comparisons

The visual comparison of the Highway dataset and the PETS 2006 dataset are shown in Fig. 5.9 and in Fig. 5.10 respectively with results based on the ground truth frame. The red bounding box indicates the occluded part, as specified in the ground truth frame. The proposed method is capable of generating both bounding boxes (represented as yellow color) on detected objects and bounding boxes on the occluded area (represented as pink color).



(a) Input frame (b) CBOT[14] (c) OIMOD[15]

(d) SAOD[37] (e) AWOD[18] (f) Proposed

**Fig. 5.9** Visual comparisons of proposed method with existing methods on the 209th frame of the Highway dataset

(a) Input frame (b) CBOT[14] (c) OIMOD[15]

(d) SAOD[37] (e) AWOD[18] (f) Proposed

**Fig. 5.2** Visual comparisons of proposed method with existing methods on the 429th frame of the PETS2006 dataset

### 5.3.2 Quantitative comparisons

The proposed method has been compared to state-of-the-art techniques using varying levels of occlusion, and it is found that the proposed method can detect an average overlapped amount of 37.84% for the Highway dataset and 39.5% for the PETS 2006 dataset when occlusion level from 20% to 40%. When the occlusion level from 40% to 70%, the proposed method can detect a 60.8% overlapped amount for the Highway dataset and 65.7% for the PETS2006 dataset. The Avg GIoU and Avg CIoU values were computed. It is observed that this value should be close to the Avg IoU value, indicating the effectiveness of the proposed method. Table 5.7 presents the results, including an average analysis of Precision, Recall, F1-Score, Intersection over Union (IoU), and Bounding Box Localization Loss (BBL) to evaluate performance under partial occlusion conditions. Our proposed method outperformed other methods in terms of F1-Score by accurately predicting bounding boxes.

**Table 5.7** Comparative performance measures under occlusion condition

| | **Highway Dataset** | | | | | **PETS2006 Dataset** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Methods** | **Pre ↑** | **Re ↑** | **F1 ↑** | **IoU ↑** | **BBL ↓** | **Pre ↑** | **Re ↑** | **F1 ↑** | **IoU ↑** | **BBL ↓** |
| CBOT[14] | 0.773 | 0.785 | 0.778 | 0.589 | 0.054 | 0.767 | 0.743 | 0.754 | 0.574 | 0.093 |
| OIMOD[15] | 0.815 | 0.804 | 0.809 | 0.625 | 0.048 | 0.742 | 0.688 | 0.713 | 0.562 | 0.131 |
| SAOD[37] | 0.863 | **0.898** | 0.880 | 0.619 | 0.051 | 0.785 | 0.726 | 0.752 | 0.594 | 0.105 |
| AWOD[18] | 0.915 | 0.881 | 0.897 | 0.644 | 0.041 | 0.816 | 0.825 | 0.820 | 0.625 | 0.072 |
| Proposed | **0.927** | 0.875 | **0.900** | **0.658** | **0.035** | **0.847** | **0.851** | **0.848** | **0.647** | **0.041** |

Bold represents good results. Red color represents second best value. ↑ High value, ↓Less value.

Pre =Precision, Re= Recall, F1 = F1-Score, IoU = Intersection over Union, BBL = Bounding Box

Through experimentation, we found that our method accurately detects objects with an occlusion rate of 65-70%. The Avg GIoU and Avg CIoU values are important for identifying the amount of overlapping, with a tighter lower bound indicating more similarity and proximity of bounding boxes. Table 5.7 illustrates the average IoU, GIoU, and CIoU values to better understand occlusion reasoning. The effectiveness of our method was evaluated through localization loss computation for different threshold values, with the proposed method only requiring the intersection area to detect occlusion relationships.

# CHAPTER 6

# CONCLUSIONS AND FUTURE SCOPE

In this chapter, the following contributions and results analysis have been summarized:

**Adaptive background subtraction and updation method:** The proposed method is based on pixel wise two level adaptive thresholding based background subtraction method to segment foreground and background pixels accurately**.** Experimental results show that the proposed method represents good precision of 0.9254 for the Highway dataset and 0.9496 for the PETS2006 dataset. It generates less False Positive Rate (FPR) than existing methods.

**Intensity ratio-based Shadow region detection and removal method:** An intensity ratio based shadow region and YCbCr color space-based shadow correction method has been proposed to remove the shadow pixels. The method has been evaluated with three Test Cases: (i) object is associated with shadow, (ii) only shadow is available and (iii) multiple shadows are available. Experimental studies reveal that:

- The proposed method is able to remove single and multiple shadows together with good shadow detection rate (92.57%) and removal rate (93.16%) in the Highway dataset under Test Case (I)
- Achieved good shadow detection rates (94.04%) and (94.57%) for Test Case (I) and Test Case (II) respectively on SBU dataset.
- Reduced false positives by correct detection of shadow pixels and their removal.

**Localizing and detecting objects under partial occlusion:** An improvement in axis aligned bounding box method has been proposed with occlusion prior condition to detect occlusion and for the degree of overlap. After experimentation, the following observations have been made:

- The proposed method can measure the occlusion ratio up to 65% and detect the occluded objects and generate less Bounding Box Localization Loss (BBL)
- The proposed method obtained good precision at 0.927 for the Highway dataset and 0.847 for the PETS2006 dataset.
- Experiments showed that the proposed can detect the objects upto 65% occlusion

## FUTURE SCOPE

There are still several possibilities for future research that could further enhance our knowledge in this area.

(i) Regeneration of occluded part

(ii) Handling heavy occlusion with multiple viewpoints

(iii) Partial occlusion detection in 3D objects

# REFERENCES


[1] T. Bouwmans, "Traditional and recent approaches in background modeling for foreground detection: An overview," *Computer Science Review*. 2014. doi: 10.1016/j.cosrev.2014.04.001.

[2] "Changedetection.net." http://jacarini.dinf.usherbrooke.ca/dataset2012

[3] J. M. McHugh, J. Konrad, V. Saligrama, and P. M. Jodoin, "Foreground-adaptive background subtraction," *IEEE Signal Process. Lett.*, 2009, doi: 10.1109/LSP.2009.2016447.

[4] R. Zhang and J. Ding, "Object tracking and detecting based on adaptive background subtraction," 2012. doi: 10.1016/j.proeng.2012.01.139.

[5] Y. Dong, T. X. Han, and G. N. DeSouza, "Illumination invariant foreground detection using multi-subspace learning," *Int. J. Knowledge-based Intell. Eng. Syst.*, vol. 14, pp. 31–41, 2010, doi: 10.3233/KES-2010-0188.

[6] W. Cao *et al.*, "Total Variation Regularized Tensor RPCA for Background Subtraction From Compressive Measurements," *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 4075–4090, 2016, doi: 10.1109/TIP.2016.2579262.

[7] Z. Liu, K. Huang, and T. Tan, "Foreground Object Detection Using Top-Down Information Based on EM Framework," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 4204–4217, 2012, doi: 10.1109/TIP.2012.2200492.

[8] T. Mahalingam and M. Subramoniam, "A robust single and multiple moving object detection, tracking and classification," *Appl. Comput. Informatics*, vol. 17, no. 1, pp. 2–18, 2017, doi: 10.1016/j.aci.2018.01.001.

[9] L. Liu and N. Sang, "Metrics for objective evaluation of background subtraction algorithms," *Proc. - 6th Int. Conf. Image Graph. ICIG 2011*, pp. 562–565, 2011, doi: 10.1109/ICIG.2011.163.

[10] S. Meshgi, K. and Ishii, "The state-of-the-art in handling occlusions for visual object tracking," *IEICE Trans. Inf. Syst.*, vol. 98, no. 7, 2015.

[11] K. Saleh, S. Szénási, and Z. Vámossy, "Occlusion Handling in Generic Object Detection: A Review," in *2021 IEEE 19th World Symposium on Applied Machine Intelligence and*

*Informatics (SAMI)*, 2021, pp. 477–484. doi: 10.1109/SAMI50585.2021.9378657.

[12] Y. Y. Li *et al.*, "occlusion," *Comput. Vis. Image Underst.*, vol. 17, no. 3, pp. 346–359, 2008, doi: https://doi.org/10.1016/j.cviu.2007.09.014.

[13] P. Shi, B. Hou, J. Chen, and Y. Zu, "An Algorithm of Occlusion Detection for the Surveillance Camera," *Sci. Program.*, vol. 2021, 2021, doi: 10.1155/2021/6698160.

[14] Z. Moutakki, I. M. Ouloul, K. Afdel, and A. Amghar, "Real-time video surveillance system for traffic management with background subtraction using codebook model and occlusion handling," *Transp. Telecommun.*, vol. 18, no. 4, pp. 297–306, 2017, doi: 10.1515/ttj-2017-0027.

[15] P. K. Mishra and G. P. Saroha, "Occlusion handling strategies for multiple moving object classification," *Int. J. Comput. Vis. Robot.*, vol. 10, no. 2, pp. 167–184, 2020, doi: 10.1504/IJCVR.2020.105683.

[16] X. Li *et al.*, "Multi-Task Structure-Aware Context Modeling for Robust Keypoint-Based Object Tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 4, pp. 915–927, 2019, doi: 10.1109/TPAMI.2018.2818132.

[17] H. A. Sulaiman *et al.*, "Distance computation using axis aligned bounding box (AABB) parallel distribution of dynamic origin point," in *2013 Annual International Conference on Emerging Research Areas and 2013 International Conference on Microelectronics, Communications and Renewable Energy*, 2013, pp. 1–6. doi: 10.1109/AICERA-ICMiCR.2013.6575955.

[18] X. Tian, H. Li, and H. Deng, "An improved object tracking algorithm based on adaptive weighted strategy and occlusion detection mechanism," *J. Algorithms Comput. Technol.*, vol. 15, 2021, doi: 10.1177/1748302620973536.

[19] H. A. Bin Sulaiman, M. A. Othman, M. Z. A. A. Aziz, and A. Bade, "Implementation of axis-aligned bounding box for opengl 3D virtual environment," *ARPN J. Eng. Appl. Sci.*, vol. 10, no. 2, pp. 701–708, 2015.

[20] E. Bochinski, V. Eiselein, and T. Sikora, "High-Speed tracking-by-detection without using image information," in *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2017, pp. 1–6. doi: 10.1109/AVSS.2017.8078516.

[21] C. Zhou and J. Yuan, "Occlusion Pattern Discovery for Object Detection and Occlusion Reasoning," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 7, pp. 2067–2080, 2020,

doi: 10.1109/TCSVT.2019.2909982.

[22] “occlusion”.

[23] B. Garcia-Garcia, T. Bouwmans, and A. J. Rosales Silva, “Background subtraction in real applications: Challenges, current models and future directions,” *Comput. Sci. Rev.*, vol. 35, p. 100204, 2020, doi: https://doi.org/10.1016/j.cosrev.2019.100204.

[24] K. P. and R. A. V., “An adaptive hybrid GMM for multiple human detection in crowd scenario,” *Multimed. Tools Appl.*, vol. 76, no. 12, pp. 14129–14149, 2017, doi: 10.1007/s11042-016-3777-4.

[25] S. Singh Sengar and S. Mukhopadhyay, “Moving object detection using statistical background subtraction in wavelet compressed domain,” *Multimed. Tools Appl.*, vol. 79, pp. 1–22, Mar. 2020, doi: 10.1007/s11042-019-08506-z.

[26] W. Li, J. Zhang, and Y. Wang, “WePBAS: A Weighted Pixel-Based Adaptive Segmenter for Change Detection,” *Sensors (Basel).*, vol. 19, no. 12, p. 2672, Jun. 2019, doi: 10.3390/s19122672.

[27] J. Liao, H. Wang, Y. Yan, and J. Zheng, “A Novel Background Subtraction Method Based on ViBe,” 2018, pp. 428–437. doi: 10.1007/978-3-319-77383-4_42.

[28] “SBU.” https://www3.cs.stonybrook.edu/~cvl/projects/shadow_noisy_label/index.html

[29] S. Murali and G. V K, “Shadow Detection and Removal from a Single Image Using LAB Color Space,” *Cybern. Inf. Technol.*, vol. 13, Mar. 2013, doi: 10.2478/cait-2013-0009.

[30] M. Wu, R. Chen, and Y. Tong, “Shadow Elimination Algorithm Using Color and Texture Features,” *Comput. Intell. Neurosci.*, vol. 2020, p. 2075781, 2020, doi: 10.1155/2020/2075781.

[31] A. Varghese and S. G, “Sample-based integrated background subtraction and shadow detection,” *IPSJ Trans. Comput. Vis. Appl.*, vol. 9, no. 1, p. 25, 2017, doi: 10.1186/s41074-017-0036-1.

[32] M. Ebner, X. Yuan, and Z. Wang, “Single-image shadow detection and removal using local colour constancy computation,” *IET Image Process.*, vol. 9, pp. 118–126, Feb. 2015, doi: 10.1049/iet-ipr.2014.0242.

[33] Y. Amit, P. Felzenszwalb, and R. Girshick, “Object Detection,” in *Computer Vision: A*

*Reference Guide*, Cham: Springer International Publishing, 2020, pp. 1–9. doi: 10.1007/978-3-030-03243-2_660-1.

[34] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, "Generalized intersection over union: A metric and a loss for bounding box regression," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2019-June, pp. 658–666, 2019, doi: 10.1109/CVPR.2019.00075.

[35] Z. Zheng, P. Wang, W. Liu, J. Li, R. Ye, and D. Ren, "Distance-IoU loss: Faster and better learning for bounding box regression," *AAAI 2020 - 34th AAAI Conf. Artif. Intell.*, no. 2, pp. 12993–13000, 2020, doi: 10.1609/aaai.v34i07.6999.

[36] B. Pepik, M. Stark, and P. Gehler, "Occlusion Patterns for Object Class Detection," *CVPR*, pp. 3286–3293, 2013, doi: 10.1109/CVPR.2013.422.

[37] Y. Yuan, J. Chu, L. Leng, J. Miao, and B.-G. Kim, "A scale-adaptive object-tracking algorithm with occlusion detection," *EURASIP J. Image Video Process.*, vol. 2020, no. 1, p. 7, 2020, doi: 10.1186/s13640-020-0496-6.

# LIST OF PAPER PUBLICATIONS

The following research papers are publishes as a part of this research work:

## Journals

1. S. Agrawal and P. Natu, "ABGS Segmenter: pixel wise adaptive background subtraction and intensity ratio based shadow removal approach for moving object detection," *Journal of Supercomputing*, Springer, 2022. **(Scopus indexed, Science Citation Index Expanded, 2.56 impact factor)**
2. S. Agrawal and P. Natu "An improved bounding box localization method to detect objects under partial occlusion", *The Visual Computer*, Springer, **(Under Review)**
3. S. Agrawal and P. Natu, "Segmentation of Moving Objects using Numerous Background Subtraction Methods for Surveillance Applications," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 3, pp. 2553–2563, 2020.

## Conference

1. S. Agrawal and P. Natu, "An improved Gaussian Mixture Method based Background Subtraction Model for Moving Object Detection in Outdoor Scene," *4th International Conference of Electrical, Computer and Communication Technology* - ICECCT 2021, IEEE Explore **(Scopus indexed)**